IN brief

1000 genomes on Amazon's cloud

Genomes free for download

The world's largest data set on human genetic variation, the 1000 Genomes Project, is now available for free through Amazon Web Services. The arrangement with Amazon, spearheaded through the US National Institutes of Health, makes data from sequencing the genomes of more than 1,700 people accessible through Amazon's cloud. Any scientist with an internet connection will be able to access all 200 terabytes of the data set in the cloud, avoiding the need to acquire additional bandwidth or data storage capacity. "It's a very large amount of data," says William Spooner, chief technology officer at Cambridge, UK–based Eagle Genomics, which provides bioinformatics services for mining genomics data. "You would need dedicated server capacity for 200 terabytes." Most pharma companies have the capacity to download that amount of data, but "these days there is no advantage to doing that," he says. Plus, working with the data in the cloud makes it easier for companies to share their analyses with university collaborators who may not have the extra server capacity, Spooner says. "What the cloud offers pharma is the ability to share analyses securely and easily with university collaborators without any risk to company firewalls." It's a good deal for both industry and public sector researchers, who will likely find the information from the 1000 Genomes Project enormously useful. But working with data in the cloud has its drawbacks. Michael Snyder, chair of genetics at Stanford University and co-founder of Personalis, a genome analysis company in Palo Alto, California, says he has used Amazon's cloud for other projects and found it difficult to get data in and out. "We started out doing stuff in the cloud and found it to be a pain," he says. Snyder says Personalis will access the data from the 1000 Genomes Project, but his team hasn't yet decided if they will do it though Amazon's cloud. Amazon Web Services could profit as well. Additional resources will be needed to crunch the data once and not everyone has this kind of computing power. Amazon can provide these resources for a fee. The 1000 Genomes Project is an international effort initiated in 2008 to collect the genomes of more than 2,600 people from 26 populations around the world (*Nat. Biotechnol.* **26**, 256, 2008). The remaining 900 samples will be sequenced this year. The move to put the project in the cloud is part of a larger effort by the US federal government to manage the deluge of 'big data' in science. The White House Office of Science and Technology Policy in March announced that more than $200 million would be doled out to six federal agencies to manage the mountains of data being created for scientific discovery. *Emily Waltz*

Table 1 Selected compounds in development for IPF

Company	Drug	Target	Status
InterMune	Esbriet	TGF-β inhibitor	Approved in Europe, but FDA has required additional phase 3 data
Boehringer Ingelheim	Vargatef	Small-molecule tyrosine kinase inhibitor against VEGFR, FGFR and PDGFR	Phase 3
Bristol-Myers Squibb	BMS-986202	Small-molecule LPA receptor-1 antagonist	Phase 2 launched; anticipated in 2012
Stromedix	STX-100	Humanized mAb targeting $\alpha_v\beta_6$	Phase 2a
Gilead	GS-6624	Humanized mAb targeting LOXL2	Phase 1

The European Union (EU) approved Esbriet in March 2011, yet InterMune failed in its first bid to attain the go-ahead from the FDA. Its prospects now hinge upon a phase 3 trial, which "will not only determine if it's approved at all in the United States," says analyst Brian Skorney of New York–based Brean Murray, Carret & Co, "but could also affect pricing in the EU." Esbriet launched in Germany last September at an annual price of about $51,000. In April, France, Spain and Italy also agreed to reimburse the treatment. If Esbriet is approved in the US, Skorney anticipates the price to hit $70,000 or more. And he also expects peak sales of $2 billion a year—with $750 million of that in Europe—if phase 3 trial results deliver. Others have published estimates that put those peak sales even higher.

According to InterMune's Welch, results from the latest Esbriet trial should come out in mid-2013. That's just about the time that phase 2a results are due for STX-100. Elsewhere, the oxindole derivative small-molecule Vargatef (BIBF 1120), made by Boehringer Ingelheim of Ingelheim am Rhein, Germany, is in phase 3 trials in Europe and Japan. This multikinase inhibitor targets the vascular endothelial growth factor receptor, fibroblast growth factor receptor and platelet-derived growth factor receptor.

That $2-billion market figure is what is driving the growing number of companies shopping for IPF treatments. In 2010, Foster City, California–based Gilead Sciences acquired Arresto Biosciences of Palo Alto, California, for $225 million, mainly to get access to GS-6624, their humanized mAb against lysyl oxidase–like-2 (LOXL2), an enzyme thought to encourage fibrosis in part through cross-linking of collagen and neovascularization in the extracellular matrix.

Likewise, New York–based Bristol-Myers Squibb (BMS) paid $325 million upfront, with up to $150 million in milestones, to buy out San Diego–based Amira Pharmaceuticals, whose pipeline included AM152, a small-molecule lysophosphatidic acid (LPA) receptor-1 antagonist that is believed to inhibit recruitment, proliferation and activation of fibroblasts. Researchers from BMS and Duke University (including Paul Noble) will be developing a phase 2 clinical protocol for the drug (now designated BMS-986202) in IPF, as well as doing biomarker validation studies. The study is expected to launch in late 2012.

At the moment, InterMune isn't betraying concern at the mounting competition (**Table 1**): "In Europe, we will have at least several years to establish Esbriet as the standard of care for appropriate patients," Welch says. He adds that there is also a growing belief that IPF, like hepatitis C, HIV and most cancers, will require multi-drug combination therapy.

The fibrosis field as a whole should benefit from the gradual uncovering of which underlying features are common to different types of fibrosis and which features distinguish particular organs. Alterations to surfactant proteins secreted by lung cells, which have a major role in pulmonary function, have been linked to some forms of IPF.

But gray areas and considerable uncertainty remain. Biogen Idec's Gilman, for example, questions whether Esbriet actually works the way it is claimed to work. "It may be a TGF-β inhibitor, but I'm not sure what the evidence is that it's a very direct inhibitor, and it may not be very effective," he says. With so many new targets and inhibitors, there are high hopes that at least some of them will work. Duke's Noble concurs: "The momentum is really going," he says. "As long as we can get the first drug approved in the US, the dam will break."

But Barrett is cautious about how many more potential $1-billion-plus compounds there are languishing in big pharma and biotech's backrooms. "There are probably some, but there aren't thousands." Atlas, he says, has at least one such project on the burner. It's a migraine drug from Lilly of Indianapolis that Atlas is funding a phase 2 trial for, hoping the pharma giant will buy back the compound on the basis of the results. "But we've looked at hundreds of such projects, and we've only found one or two that were interesting," he says.

Malorye Allison *Acton, Massachusetts*